IEEE Xplore RELEASE 2.1

Welcome United States Patent and Trademark Office

Results for "(((natural and language and term and relevance)<in>metadata)) <and> (pyr >= 1950 <..."

e-mail printer friendly

Your search matched **0** documents.

A maximum of **100** results are displayed, **25** to a page, sorted by **Relevance** in **Descending** order.

» Search Options

View Session History

New Search

Modify Search

(((natural and language and term and relevance)<in>metadata)) <and> (pyr >= 1950 Search

Check to search only within this results set

Display Format: Citation Citation & Abstract

» Key

IEEE JNL	IEEE Journal or Magazine
IEE JNL	IEE Journal or Magazine
IEEE CNF	IEEE Conference Proceeding
IEE CNF	IEE Conference Proceeding
IEEE STD	IEEE Standard

No results were found.

Please edit your search criteria and try again. Refer to the Help pages if you need assistance revising your search.

Indexed by Inspec

Results for "((natural and language and term and relevance)<in>metadata)"

e-mail printer friendly

Your search matched 6 of 1415139 documents.

A maximum of 100 results are displayed, 25 to a page, sorted by Relevance in Descending order.

» Search Options

View Session History

New Search

Modify Search

((natural and language and term and relevance)<in>metadata) Search

Check to search only within this results set

» Key

IEEE JNL	IEEE Journal or Magazine
IEE JNL	IEE Journal or Magazine
IEEE CNF	IEEE Conference Proceeding
IEE CNF	IEE Conference Proceeding
IEEE STD	IEEE Standard

Display Format: Citation Citation & Abstract

view selected items Select All Deselect All

1. **Chinese query expansion based on related term group**
Tingting He; Xinhui Tu; Guozhong Qu;
Natural Language Processing and Knowledge Engineering, 2005. IEEE NLP-KE '05. Proceedings of 2005 IEEE International Conference on
30 Oct.-1 Nov. 2005 Page(s):483 - 487
Digital Object Identifier 10.1109/NLPKE.2005.1598785
AbstractPlus | Full Text: PDF(976 KB) IEEE CNF
Rights and Permissions

2. **Query expansion based on term similarity tree model**
Qianli Jin; Jun Zhao; Bo Xu;
Natural Language Processing and Knowledge Engineering, 2003. Proceedings. 2003 International Conference on
26-29 Oct. 2003 Page(s):400 - 406
Digital Object Identifier 10.1109/NLPKE.2003.1275938
AbstractPlus | Full Text: PDF(400 KB) IEEE CNF
Rights and Permissions

3. **Combining words and object-based visual features in image retrieval**
Nakagawa, A.; Kutics, A.; Tanaka, K.; Nakajima, M.;
Image Analysis and Processing, 2003.Proceedings. 12th International Conference on
17-19 Sept. 2003 Page(s):354 - 359
Digital Object Identifier 10.1109/ICIAP.2003.1234075
AbstractPlus | Full Text: PDF(1597 KB) IEEE CNF
Rights and Permissions

4. **A fuzzy set approach to extracting keywords from abstracts**
Makrehchi, M.; Kamel, M.;
Fuzzy Information, 2004. Processing NAFIPS '04. IEEE Annual Meeting of the
Volume 2, 27-30 June 2004 Page(s):528 - 532 Vol.2
Digital Object Identifier 10.1109/NAFIPS.2004.1337356
AbstractPlus | Full Text: PDF(507 KB) IEEE CNF
Rights and Permissions

5. **Ontology-based medical image annotation with description logics**
Bo Hu; Dasmahapatra, S.; Lewis, P.; Shadbolt, N.;
Tools with Artificial Intelligence, 2003. Proceedings. 15th IEEE International Conference on
3-5 Nov. 2003 Page(s):77 - 82
AbstractPlus | Full Text: PDF(386 KB) IEEE CNF
Rights and Permissions

6. **Mining answers in German Web pages**
Neumann, G.; Feiyu Xu;
Web Intelligence, 2003. WI 2003. Proceedings. IEEE/WIC International Conference on
13-17 Oct. 2003 Page(s):125 - 131

AbstractPlus | Full Text: PDF(257 KB) **IEEE CNF**
Rights and Permissions

Google Scholar BETA

natural language term overall degree of releva Search

Advanced Scholar Search
Scholar Preferences
Scholar Help

"of" is a very common word and was not included in your search. [details]

Scholar **All articles** Recent articles Results **1 - 10** of about **15** for **natural language term overall degree of relevance ratin**

Natural Language Information Retrieval: TREC-5 Report - group of 13 »
T Strzalkowski, JP Carballo - Text REtrieval Conference, 1998 - ai.mit.edu
... this paper we describe the **overall** organization of ... **natural** and processing+**language** from "**natural language** processing" as ... up about 8% of all **terms** generated ...
Cited by 125 - Related Articles - View as HTML - Web Search - BL Direct

Information retrieval using robust **natural language** processing - group of 8 »
T Strzalkowski, B Vauthey - Proceedings of the 30th conference on Association for ..., 1992 - portal.acm.org
... that the similarities are calculated using **term** co ... appropriate to predict similarity between **language** and logarithm ... of their co-occurrence with **naturaL** to This ...
Cited by 37 - Related Articles - Web Search

Natural language information retrieval: progress report - group of 4 »
J Perez-Carballo, **T Strzalkowski** - Information Processing and Management, 2000 - alizarin.njit.edu
... and proper names computed using **Natural Language** Processing techniques ... weighting scheme for compound **terms**, including phrases ... names, leads to an **overall** gain in ...
Cited by 24 - Related Articles - View as HTML - Web Search

Evaluating **natural language** processing techniques in information retrieval
T Strzalkowski, F Lin, J Wang, J Perez-Carballo - **Natural Language** Information Retrieval. Kluwer, Boston, MA, 1999 - nlp.korea.ac.kr
... **Overall** 7% precision gain (gain(13): 754%, loss(18 ... Using linguistic **terms**, such as phrases, head ... [1] Tomek Strzalkowski, "**Natural Language** Information Retrieval ...
Cited by 22 - Related Articles - View as HTML - Web Search

Building effective queries in **natural language** information retrieval - group of 9 »
T Strzalkowski, F Lin, J Perez-Carballo, J Wang - ... of the fifth conference on Applied **natural language** ..., 1997 - acl.ldc.upenn.edu
... of this paper we describe the **overall** organi- zation ... a suite of advanced **natural language** processing techniques ... The following **term** extraction methods have been ...
Cited by 10 - Related Articles - View as HTML - Web Search

Summarization-based Query Expansion in Information Retrieval - group of 10 »
T Strzalkowski, J Wang, B Wise - Proceedings of the 36th conference on Association for ..., 1998 - acl.ldc.upenn.edu
... An alternative to **term**-only expansion is a ... These documents, irrespective of their **overall** relevancy to ... through a series of **natural language** processing steps ...
Cited by 12 - Related Articles - View as HTML - Web Search - BL Direct

Natural language information retrieval in digital libraries - group of 3 »
T Strzalkowski, J Perez-Carballo, M Marinescu - Proceedings of the first ACM international conference on ..., 1996 - portal.acm.org
... a query already containing "**natural language**" because "**natural** ... **natural** number", "**natural** logarithm", "**natural** ... the compound **term** "illegal activity ...
Cited by 6 - Related Articles - Web Search

[PS] Towards the next generation information retrieval - group of 4 »
T Strzalkowski, GC Stein, GB Wise, A Bagga - 6eme Conference de Recherche d'Information Assistee par ..., 2000 - ai.mit.edu
... These documents, irrespective of their **overall** relevancy to ... of scores because of possible **term** repetitions. ... 1. The initial **natural language** topic statement is ...
Cited by 7 - Related Articles - View as HTML - Web Search

Natural language information retrieval: TIPSTER-2 final report - group of 3 »
T Strzalkowski - Proceedings of a workshop on held at Vienna, Virginia: May 6 ..., 1996 - portal.acm.org
... **Overall**, our system performed quite well as our posi ... in each case determining a **degree** of **relevance** ... For example, the **term natural language** may be considered to ...
Cited by 2 - Related Articles - Web Search

Cross-document summarization by concept classification - group of 4 »
H Hardy, N Shimizu, **T Strzalkowski**, L Ting, X ... - Proceedings of the 25th annual international ACM SIGIR ..., 2002 - portal.acm.org
... sentences by combining a unigram **language** model approach ... which usually correspond to the **natural** paragraphs designed by ... be at least two common **terms** among the ...
Cited by 10 - Related Articles - Web Search

Google ▶
Result Page: 1 2 **Next**

Terms used **degree of relevance** and **database** and **query** Found **35,542** of **185,178**

Sort results by: relevance
Display results: expanded form
Save results to a Binder
Search Tips
Open results in a new window

Try an Advanced Search
Try this search in The ACM Guide

Results 21 - 40 of 200 Result page: previous 1 **2** 3 4 5 6 7 8 9 10 next
Best 200 shown Relevance scale

21 Efficient and effective metasearch for a large number of text databases
Clement Yu, Weiyi Meng, King-Lup Liu, Wensheng Wu, Naphtali Rishe
November 1999 **Proceedings of the eighth international conference on Information and knowledge management**
Publisher: ACM Press
Full text available: pdf(1.04 MB) Additional Information: full citation, abstract, references, citings, index terms

Metasearch engines can be used to facilitate ordinary users for retrieving information from multiple local sources (text databases). In a metasearch engine, the contents of each local database is represented by a representative. Each user query is evaluated against the set of representatives of all databases in order to determine the appropriate databases to search. When the number of databases is very large, say in the order of tens of thousands or more, then a traditional metasearch engin ...

22 Evaluating database selection techniques: a testbed and experiment
James C. French, Allison L. Powell, Charles L. Viles, Travis Emmitt, Kevin J. Prey
August 1998 **Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval**
Publisher: ACM Press
Full text available: pdf(1.01 MB) Additional Information: full citation, references, citings, index terms

23 The OODB path-method generator (PMG) using access weights and precomputed access relevance
Ashish Mehta, James Geller, Yehoshua Perl, Erich Neuhold
February 1998 **The VLDB Journal — The International Journal on Very Large Data Bases**, Volume 7 Issue 1
Publisher: Springer-Verlag New York, Inc.
Full text available: pdf(265.48 KB) Additional Information: full citation, abstract, citings, index terms

A *path-method* is used as a mechanism in object-oriented databases (OODBs) to retrieve or to update information relevant to one class that is not stored with that class but with some other class. A path-method is a method which traverses from one class through a chain of connections between classes and accesses information at another class. However, it is a difficult task for a casual user or even an application programmer to write path-methods to facilitate queries. This is because it mig ...

Keywords: Access relevance, Access weight, OODB queries, Object-oriented databases, Path-method, Traversal algorithms

24 Enhancing relevance feedback in image retrieval using unlabeled data

Zhi-Hua Zhou, Ke-Jia Chen, Hong-Bin Dai
April 2006 **ACM Transactions on Information Systems (TOIS)**, Volume 24 Issue 2
Publisher: ACM Press
Full text available: pdf(1.23 MB) Additional Information: full citation, abstract, references, index terms

Relevance feedback is an effective scheme bridging the gap between high-level semantics and low-level features in content-based image retrieval (CBIR). In contrast to previous methods which rely on labeled images provided by the user, this article attempts to enhance the performance of relevance feedback by exploiting unlabeled images existing in the database. Concretely, this article integrates the merits of semisupervised learning and active learning into the relevance feedback process. In det ...

Keywords: Relevance feedback, active learning, content-based image retrieval machine learning, learning with unlabeled data, semisupervised learning

25 Term relevance feedback and query expansion: relation to design

Amanda Spink
August 1994 **Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval**
Publisher: Springer-Verlag New York, Inc.
Full text available: pdf(686.20 KB) Additional Information: full citation, references, citings, index terms

26 Appication models: Client-system collaboration for legal corpus selection in an online production environment

Jack G. Conrad, Joanne R. S. Claussen
June 2003 **Proceedings of the 9th international conference on Artificial intelligence and law**
Publisher: ACM Press
Full text available: pdf(239.10 KB) Additional Information: full citation, abstract, references

The continued growth of very large data environments such as Westlaw and Dialog, in addition to the World Wide Web, increases the importance of effective and efficient database selection and searching. Current research focuses largely on completely autonomous and automatic selection, searching, and results merging in distributed environments. This fully automatic approach has significant deficiencies, including reliance upon thresholds below which databases with relevant documents are not search ...

Keywords: database selection, query categorization, user interaction

27 Annotated Bibliography Relating to Automatic Indexing in Information Retrieval

September 1986 **ACM SIGIR Forum**, Volume 21 Issue 1-2
Publisher: ACM Press
Full text available: pdf(938.18 KB) Additional Information: full citation

28 Information access and retrieval (IAR): Relevance feedback methods for logo and trademark image retrieval on the web

Euripides G. M. Petrakis, Klaydios Kontis, Epimenidis Voutsakis, Evangelos E. Milios
April 2006 **Proceedings of the 2006 ACM symposium on Applied computing SAC '06**
Publisher: ACM Press
Full text available: pdf(110.23 KB) Additional Information: full citation, abstract, references, index terms

Relevance feedback is the state-of-the-art approach for adjusting query results to the needs of the users. This work extends the existing framework of image retrieval with relevance feedback on the Web by incorporating text and image content into the search and feedback process. Some of the most powerful relevance feedback methods are implemented and tested on a fully automated Web retrieval system with more than 250,000 logo and trademark images. This evaluation demonstrates that term re-weight ...

Keywords: image retrieval, relevance feedback, world wide web

29 Querying and web: Efficient query processing in geographic web search engines

Yen-Yu Chen, Torsten Suel, Alexander Markowetz

June 2006 **Proceedings of the 2006 ACM SIGMOD international conference on Management of data SIGMOD '06**

Publisher: ACM Press

Full text available: pdf(296.76 KB) Additional Information: full citation, abstract, references

Geographic web search engines allow users to constrain and order search results in an intuitive manner by focusing a query on a particular geographic region. Geographic search technology, also called *local search*, has recently received significant interest from major search engine companies. Academic research in this area has focused primarily on techniques for extracting geographic knowledge from the web. In this paper, we study the problem of efficient query processing in scalable geogr ...

30 Technical session 1: content-based image retrieval: A novel log-based relevance feedback technique in content-based image retrieval

Chu-Hong Hoi, Michael R. Lyu

October 2004 **Proceedings of the 12th annual ACM international conference on Multimedia**

Publisher: ACM Press

Full text available: pdf(228.62 KB) Additional Information: full citation, abstract, references, index terms

Relevance feedback has been proposed as an important technique to boost the retrieval performance in content-based image retrieval (CBIR). However, since there exists a semantic gap between low-level features and high-level semantic concepts in CBIR, typical relevance feedback techniques need to perform a lot of rounds of feedback for achieving satisfactory results. These procedures are time-consuming and may make the users bored in the retrieval tasks. For a long-term study purpose in CBIR, ...

Keywords: content-based image retrieval, relevance feedback, support vector machines, users logs

31 Comparing the performance of collection selection algorithms

Allison L. Powell, James C. French

October 2003 **ACM Transactions on Information Systems (TOIS)**, Volume 21 Issue 4

Publisher: ACM Press

Full text available: pdf(668.40 KB) Additional Information: full citation, abstract, references, citings, index terms

The proliferation of online information resources increases the importance of effective and efficient information retrieval in a multicollection environment. Multicollection searching is cast in three parts: collection selection (also referred to as database selection), query processing and results merging. In this work, we focus our attention on the evaluation of the first step, collection selection.In this article, we present a detailed discussion of the methodology that we used to evaluate an ...

Keywords: Collection selection, database selection, distributed information retrieval, distributed text retrieval, metasearch engine, resource discovery, resource ranking, resource selection, server ranking, server selection, text retrieval

32 On the measurement of inter-linker consistency and retrieval effectiveness in hypertext databases
David Ellis, Jonathan Furner-Hines, Peter Willett
August 1994 **Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval**
Publisher: Springer-Verlag New York, Inc.
Full text available: pdf(1.04 MB) Additional Information: full citation, references, citings, index terms, review

33 LyberWorld—a visualization user interface supporting fulltext retrieval
Matthias Hemmje, Clemens Kunkel, Alexander Willett
August 1994 **Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval**
Publisher: Springer-Verlag New York, Inc.
Full text available: pdf(1.31 MB) Additional Information: full citation, references, citings, index terms

34 Querying and web: To search or to crawl?: towards a query optimizer for text-centric tasks
Panagiotis G. Ipeirotis, Eugene Agichtein, Pranay Jain, Luis Gravano
June 2006 **Proceedings of the 2006 ACM SIGMOD international conference on Management of data SIGMOD '06**
Publisher: ACM Press
Full text available: pdf(625.16 KB) Additional Information: full citation, abstract, references, index terms

Text is ubiquitous and, not surprisingly, many important applications rely on textual data for a variety of tasks. As a notable example, information extraction applications derive structured relations from unstructured text; as another example, focused crawlers explore the web to locate pages about specific topics. Execution plans for text-centric tasks follow two general paradigms for processing a text database: either we can scan, or 'crawl," the text database or, alternatively, we can exploit ...

Keywords: focused crawling, information extraction, metasearching, query optimization, research, text databases

35 Query enhancement by user profiles
Robert R. Korfhage
July 1984 **Proceedings of the 7th annual international ACM SIGIR conference on Research and development in information retrieval**
Publisher: British Computer Society
Full text available: pdf(519.18 KB) Additional Information: full citation, abstract, references, citings

We describe a theoretical model and an on-going series of experiments aimed at a priori query enhancement. The model presents a synthesis of concepts from retrospective and current awareness retrieval systems, employing the user profile as a factor in interpreting a query. It is expected that this will provide a more personalized response to queries.

36 Building effective queries in natural language information retrieval
Tomek Strzalkowski, Fang Lin, Jose Perez-Carballo, Jin Wang
March 1997 **Proceedings of the fifth conference on Applied natural language processing**
Publisher: Morgan Kaufmann Publishers Inc.
Full text available: pdf(771.03 KB) Additional Information: full citation, abstract, references, citings

Publisher Site

In this paper we report on our natural language information retrieval (NLIR) project as related to the recently concluded 5th Text Retrieval Conference (TREC-5). The main thrust of this project is to use natural language processing techniques to enhance the effectiveness of full-text document retrieval. One of our goals was to demonstrate that robust if relatively shallow NLP can help to derive a better representation of text documents for statistical search. Recently, we have turned our attenti ...

37 A highly scalable and effective method for metasearch

Weiyi Meng, Zonghuan Wu, Clement Yu, Zhuogang Li

July 2001 **ACM Transactions on Information Systems (TOIS)**, Volume 19 Issue 3

Publisher: ACM Press

Full text available: pdf(653.63 KB) Additional Information: full citation, abstract, references, citings, index terms

A metasearch engine is a system that supports unified access to multiple local search engines. Database selection is one of the main challenges in building a large-scale metasearch engine. The problem is to efficiently and accurately determine a small number of potentially useful local search engines to invoke for each user query. In order to enable accurate selection, metadata that reflect the contents of each search engine need to be collected and used. This article proposes a highly scalable ...

Keywords: Database selection, distributed text retrieval, metasearch engine, resource discovery

38 The SIFT information dissemination system

Tak W. Yan, Hector Garcia-Molina

December 1999 **ACM Transactions on Database Systems (TODS)**, Volume 24 Issue 4

Publisher: ACM Press

Full text available: pdf(220.77 KB) Additional Information: full citation, abstract, references, citings, index terms

Information dissemination is a powerful mechanism for finding information in wide-area environments. An information dissemination server accepts long-term user queries, collects new documents from information sources, matches the documents against the queries, and continuously updates the users with relevant information. This paper is a retrospective of the Stanford Information Filtering Service (SIFT), a system that as of April 1996 was processing over 40,000 worldwide subscriptions and ov ...

Keywords: Boolean queries, dissemination, filtering, indexing, vector space queries

39 Indexing music and Chinese text: Looking for new, not known music only: music retrieval by melody style

Fang-Fei Kuo, Man-Kwan Shan

June 2004 **Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries**

Publisher: ACM Press

Full text available: pdf(526.55 KB) Additional Information: full citation, abstract, references, index terms

With the growth of digital music, content-based music retrieval (CBMR) has attracted increasingly attention. For most CBMR systems, the task is to return music objects similar to query in syntactic properties such as pitch and interval contour sequence. These approaches provide users the capability to look for music that has been heard. However, sometimes, listeners are looking, not for music they have been known, but for music that is new to them. Moreover, people sometimes want to retrieve mus ...

Keywords: content-based music retrieval, music classification, music style mining, query by melody style

40 Dynamic query interpretation in relational databases

A. D'Atri, P. Di Felice, M. Moscarini

June 1987 **Proceedings of the sixth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems**

Publisher: ACM Press

Full text available: pdf(693.86 KB) Additional Information: full citation, abstract, references, index terms

A new dynamic approach to the problem of determining the correct interpretation of a logically independent query to a relational database is described. The proposed disambiguating process is based on a simple user-system dialogue that consists in a sequence of decisions about the relevance (or not) of an attribute with respect to the user interpretation

Results 21 - 40 of 200 Result page: previous 1 **2** 3 4 5 6 7 8 9 10 next

Terms used **natural** and **language** and **term** and **relavance** Found **92,939** of **185,178**

Sort results by: relevance
Display results: expanded form
Save results to a Binder
Search Tips
Open results in a new window

Try an Advanced Search
Try this search in The ACM Guide

Results 1 - 20 of 200 Result page: **1** 2 3 4 5 6 7 8 9 10 next
Best 200 shown Relevance scale

1 Some Comments On EQS, A Near Term Natural Language Data Base Query System
William A. Martin
December 1978 **Proceedings of the 1978 annual conference**
Publisher: ACM Press
Full text available: pdf(665.01 KB) Additional Information: full citation, abstract, references, citings, index terms

Problems and possibilities for near term natural language query systems are discussed, with emphasis on the author's own system, EQS. First, the general objectives for near term systems in the areas of syntax, world knowledge, discourse, and problem solving are considered. Next, a comparison is made between the ATN parsing strategies in LADDER, ROBOT, PLANES, and EQS. Evidence for the importance of giving answers to queries not directly available in the data base is given together with some ...

Keywords: Data base query, Natural language, Semantic data models, Semantic networks

2 Natural language querying of historical databases
James Clifford
December 1988 **Computational Linguistics**, Volume 14 Issue 4
Publisher: MIT Press
Full text available: pdf(2.82 MB) Additional Information: full citation, abstract, references, citings
Publisher Site

In this paper we examine the connection between two areas of semantics, namely the semantics of historical databases and the semantics of natural language querying, and link them together via a common view of the semantics of time. Since the target application domain is an historical database, we present the essential features of the Historical Relational Database Model (HRDM), an extension to the relational model motivated by the desire to incorporate more "real world" semantics into a database ...

3 Reasoning about inconsistencies in natural language requirements
Vincenzo Gervasi, Didar Zowghi
July 2005 **ACM Transactions on Software Engineering and Methodology (TOSEM)**, Volume 14 Issue 3
Publisher: ACM Press
Full text available: pdf(2.76 MB) Additional Information: full citation, abstract, references, index terms

The use of logic in identifying and analyzing inconsistency in requirements from multiple stakeholders has been found to be effective in a number of studies. Nonmonotonic logic is

a theoretically well-founded formalism that is especially suited for supporting the evolution of requirements. However, direct use of logic for expressing requirements and discussing them with stakeholders poses serious usability problems, since in most cases stakeholders cannot be expected to be fluent with formal log ...

Keywords: Requirements, default logic, inconsistency, natural language

4 Non-singular concepts in natural language discourse
Tomek Strzalkowski, Nick Cercone
September 1989 **Computational Linguistics**, Volume 15 Issue 3
Publisher: MIT Press
Full text available: pdf(1.86 MB) Additional Information: full citation, abstract, references
Publisher Site

We introduce a new approach to representing and manipulating various types of non-singular concepts in natural language discourse. The representation we describe is based on a partially ordered structure of levels in which the objects of the same relative singularity are assigned to the same level. Our choice of the representation has been motivated by the following main concerns: 1. The representation should systematically distinguish between those language terms that are used to refer to objec ...

5 Metric details for natural-language spatial relations
Max J. Egenhofer, A. Rashid B. M. Shariff
October 1998 **ACM Transactions on Information Systems (TOIS)**, Volume 16 Issue 4
Publisher: ACM Press
Full text available: pdf(1.47 MB) Additional Information: full citation, abstract, references, citings, index terms

Spatial relations often are desired answers that a geographic information system (GIS) should generate in response to a user's query. Current GIS's provide only rudimentary support for processing and interpreting natural-language-like spatial relations, because their models and representations are primarily quantitative, while natural-language spatial relations are usually dominated by qualitative properties. Studies of the use of spatial relations in natural language showed that topology ...

Keywords: GIS, Metric refinements, geographic information systems, spatial relations, topological relations

6 Natural language processing for information retrieval
David D. Lewis, Karen Spärck Jones
January 1996 **Communications of the ACM**, Volume 39 Issue 1
Publisher: ACM Press
Full text available: pdf(602.45 KB) Additional Information: full citation, references, citings, index terms

7 Session V: PHRAN: a knowledge-based natural language understander
Robert Wilensky, Yigal Arens
June 1980 **Proceedings of the 18th annual meeting on Association for Computational Linguistics**
Publisher: Association for Computational Linguistics
Full text available: pdf(578.87 KB) Additional Information: full citation, abstract, references, citings
Publisher Site

We have developed an approach to natural language processing in which the natural language processor is viewed as a knowledge-based system whose knowledge is about the meanings of the utterances of its language. The approach is oriented around the phrase

rather than the word as the basic unit. We believe that this paradigm for language processing not only extends the capabilities of other natural language systems, but handles those tasks that previous systems could perform in a more systematic a ...

8 Natural language information retrieval in digital libraries

Tomek Strzalkowski, Jose Perez-Carballo, Mihnea Marinescu
April 1996 **Proceedings of the first ACM international conference on Digital libraries**
Publisher: ACM Press
Full text available: pdf(1.03 MB) Additional Information: full citation, references, index terms

9 A direct manipulation interface for boolean information retrieval via natural language query

P. G. Anick, J. D. Brennan, R. A. Flynn, D. R. Hanssen, B. Alvey, J. M. Robbins
December 1989 **Proceedings of the 13th annual international ACM SIGIR conference on Research and development in information retrieval**
Publisher: ACM Press
Full text available: pdf(1.39 MB) Additional Information: full citation, abstract, references, citings, index terms

This paper describes the design of a direct manipulation user interface for Boolean information retrieval. Intended to overcome the difficulties of manipulating explicit Boolean queries as well as the "black box" drawbacks of so-called natural language query systems, the interface presents a two-dimensional graphical representation of a user's natural language query which not only exposes heuristic query transformations performed by the system, but also supports query reformulat ...

10 IR-7 (information retrieval): natural language processing for IR: Distributional term representations: an experimental comparison

Alberto Lavelli, Fabrizio Sebastiani, Roberto Zanoli
November 2004 **Proceedings of the thirteenth ACM international conference on Information and knowledge management CIKM '04**
Publisher: ACM Press
Full text available: pdf(185.23 KB) Additional Information: full citation, abstract, references, index terms

A number of content management tasks, including term categorization, term clustering, and automated thesaurus generation, view natural language terms (e.g. words, noun phrases) as first-class objects, i.e. as objects endowed with an internal representation which makes them suitable for explicit manipulation by the corresponding algorithms. The information retrieval (IR) literature has traditionally used an extensional (aka distributional) representation for terms ...

Keywords: term classification, term clustering, term representation

11 Designing a Portable Natural Language Database Query System

S. Jerrold Kaplan
March 1984 **ACM Transactions on Database Systems (TODS)**, Volume 9 Issue 1
Publisher: ACM Press
Full text available: pdf(1.55 MB) Additional Information: full citation, abstract, references, citings, index terms

One barrier to the acceptance of natural language database query systems is the substantial installation effort required for each new database. Much of this effort involves the encoding of semantic knowledge for the domain of discourse, necessary to correctly interpret and respond to natural language questions. For such systems to be practical, techniques must be developed to increase their portability to new domains. This paper discusses several issues involving the portability ...

12 Information retrieval using robust natural language processing

Tomek Strzalkowski, Barbara Vauthey

June 1992 **Proceedings of the 30th annual meeting on Association for Computational Linguistics**

Publisher: Association for Computational Linguistics

Full text available: pdf(772.67 KB) Publisher Site Additional Information: full citation, abstract, references, citings

We developed a prototype information retrieval system which uses advanced natural language processing techniques to enhance the effectiveness of traditional key-word based document retrieval. The backbone of our system is a statistical retrieval engine which performs automated indexing of documents, then search and ranking in response to user queries. This core architecture is augmented with advanced natural language processing tools which are both robust and efficient. In early experiments, the ...

13 An intelligent approach to handling imperfect information in concept-based natural language queries

Vesper Owei

July 2002 **ACM Transactions on Information Systems (TOIS)**, Volume 20 Issue 3

Publisher: ACM Press

Full text available: pdf(5.44 MB) Additional Information: full citation, abstract, references, index terms

Missing information, imprecision, inconsistency, vagueness, uncertainty, and ignorance abound in information systems. Such imperfection is a fact of life in database systems. Although these problems are widely studied in relational database systems, this is not the case in conceptual query systems. And yet, concept-based query languages have been proposed and some are already commercial products. It is therefore imperative to study these problems in concept-based query languages, with a view to ...

Keywords: ambiguous query, anaphoric query, concept-based query, conceptual query language, elliptical query, imperfect queries, incomplete information, inconsistency, inexplicit query, missing information, natural language interface, natural language query, semantically mismatched query

14 Special issue on user modeling: Modeling the user in natural language systems

Robert Kass, Tim Finin

September 1988 **Computational Linguistics**, Volume 14 Issue 3

Publisher: MIT Press

Full text available: pdf(2.13 MB) Publisher Site Additional Information: full citation, abstract, references, citings

For intelligent interactive systems to communicate with humans in a natural manner, they must have knowledge about the system users. This paper explores the role of *user modeling* in such systems. It begins with a characterization of what a user model is and how it can be used. The types of information that a user model may be required to keep about a user are then identified and discussed. User models themselves can vary greatly depending on the requirements of the situation and the imple ...

15 Research in information extraction & document detection: Natural language information retrieval: TIPSTER-2 final report

Tomek Strzalkowski

May 1996 **Proceedings of a workshop on held at Vienna, Virginia: May 6-8, 1996**

Publisher: Association for Computational Linguistics

Full text available: pdf(561.17 KB) Additional Information: full citation, abstract, references

We report on the joint GE/NYU natural language information retrieval project as related to the Tipster Phase 2 research conducted initially at NYU and subsequently at GE R&D Center and NYU. The evaluation results discussed here were obtained in connection with

the 3rd and 4th Text Retrieval Conferences (TREC-3 and TREC-4). The main thrust of this project is to use natural language processing techniques to enhance the effectiveness of full-text document retrieval. During the course of the four TR ...

16 Linguistics: linguistic topics: Levels of representation in natural language based information systems and their relation to the methodology of computational linguistics
G. Zifonun
September 1980 **Proceedings of the 8th conference on Computational linguistics**
Publisher: Association for Computational Linguistics
Full text available: pdf(577.84 KB) Additional Information: full citation, references

17 Natural language question-answering systems: 1969
Robert F. Simmons
January 1970 **Communications of the ACM**, Volume 13 Issue 1
Publisher: ACM Press
Full text available: pdf(2.15 MB) Additional Information: full citation, abstract, references, citings

Recent experiments in programming natural language question-answering systems are reviewed to summarize the methods that have been developed for syntactic, semantic, and logical analysis of English strings. It is concluded that at least minimally effective techniques have been devised for answering questions from natural language subsets in small scale experimental systems and that a useful paradigm has evolved to guide research efforts in the field. Current approaches to semantic analysis ...

Keywords: artificial intelligence, fact retrieval, language processing, natural language, question-answering system, semantics

18 Research actvities on natural language processing of the FGCS project
Toshio Yokoi, Kuniaki Mukai, Hideo Miyoshi, Yuichi Tanaka
November 1986 **Proceedings of 1986 ACM Fall joint computer conference**
Publisher: IEEE Computer Society Press
Full text available: pdf(579.38 KB) Additional Information: full citation, references, index terms

19 Unsupervised learning of the morphology of a natural language
John Goldsmith
June 2001 **Computational Linguistics**, Volume 27 Issue 2
Publisher: MIT Press
Full text available: pdf(3.19 MB) Additional Information: full citation, abstract, references, citings
Publisher Site

This study reports the results of using minimum description length (MDL) analysis to model unsupervised learning of the morphological segmentation of European languages, using corpora ranging in size from 5,000 words to 500,000 words. We develop a set of heuristics that rapidly develop a probabilistic morphological grammar, and use MDL as our primary tool to determine whether the modifications proposed by the heuristics will be adopted or not. The resulting grammar matches well the analysis that ...

20 Summarizing natural language database responses
Jugal K. Kalita, Marlene L. Jones, Gordon I. McCalla
April 1986 **Computational Linguistics**, Volume 12 Issue 2
Publisher: MIT Press
Full text available: pdf(2.04 MB) Additional Information: full citation, abstract, references, citings

Publisher Site

In a human dialogue it is usually considered inappropriate if one conversant monopolizes the conversation. Similarly it can be inappropriate for a natural language database interface to respond with a lengthy list of data. A non-enumerative "summary" response is less verbose and often avoids misleading the user where an extensional response might.In this paper we investigate the problem of generating such discourse-oriented concise responses. We present details of the design and implementation o ...

Results 1 - 20 of 200 Result page: **1** 2 3 4 5 6 7 8 9 10 next

Design and evaluation of algorithms for image retrieval by spatial similarity

Full text Pdf (1.89 MB)

Source **ACM Transactions on Information Systems (TOIS)** archive
Volume 13 , Issue 2 (April 1995) table of contents
Pages: 115 - 144
Year of Publication: 1995
ISSN:1046-8188

Authors Venkat N. Gudivada Ohio Univ., Athens
Vijay V. Raghavan Univ. of Southwest Louisiana, Lafayette

Publisher ACM Press New York, NY, USA

↑ ABSTRACT

Similarity-based retrieval of images is an important task in many image database applications. A major class of users' requests requires retrieving those images in the database that are spatially similar to the query image. We propose an algorithm for computing the spatial similarity between two symbolic images. A symbolic image is a logical representation of the original image where the image objects are uniquely labeled with symbolic names. Spatial relationships in a symbolic image are represented as edges in a weighted graph referred to as spatial-orientation graph. Spatial similarity is then quantified in terms of the number of, as well as the extent to which, the edges of the spatial-orientation graph of the database image conform to the corresponding edges of the spatial-orientation graph of the query image.The proposed algorithm is robust in the sense that it can deal with translation, scale, and rotational variances in images. The algorithm has quadratic time complexity in terms of the total number of objects in both the database and query images. We also introduce the idea of quantifying a system's retrieval quality by having an expert specify the expected rank ordering with respect to each query for a set of test queries. This enables us to assess the quality of algorithms comprehensively for retrieval in image databases. The characteristics of the proposed algorithm are compared with those of the previously available algorithms using a testbed of images. The comparison demonstrated that our algorithm is not only more efficient but also provides a rank ordering of images that consistently matches with the expert's expected rank ordering.

↑ REFERENCES

Note: OCR errors may be found in this Reference List extracted from the full text article. ACM has opted to expose the complete List rather than only correct and linked references.

1 P. Bollmann , F. Jochum , U. Reiner , V. Weissmann , H. Zuse, The LIVE-project: retrieval experiments based on evaluation viewpoints, Proceedings of the 8th annual international ACM SIGIR conference on Research and development in information retrieval, p.213-214, June 05-07, 1985, Montreal, Quebec, Canada

2 C. C. Chang , S. Y. Lee, Retrieval of similar pictures on pictorial databases, Pattern Recognition, v.24 n.7, p.675-681, 1991

3 Margaret Irvine Chock, A data base management system for image processing, 1982

4 FISHER, W. 1958. On grouping for maximum homogeneity. J. Am. Stat. Assoc. 53, 789-798.

5 James D. Foley , Andries van Dam , Steven K. Feiner , John F. Hughes, Computer graphics: principles and practice (2nd ed.), Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 1990

6 Michael R. Garey , David S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness, W. H. Freeman & Co., New York, NY, 1979

7 GUDWADA, V.N. 1993. A unified framework for retrieval in image databases. Ph.D. dissertation, Univ. of Southwestern Louisiana, Lafayette, La.

8 Suh-Yin Lee , Man-Kwan Shan , Wei-Pang Yang, Similarity retrieval of iconic image database, Pattern Recognition, v.22 n.6, p.675-682, November 1989

9 RAGHAVAN, V. V. AND GUDIVADA, V. N. 1990. A domain independent similarity measure for symbolic images. In 1st Indian Computing Congress (Hyderabad, India, Nov.). 195-203.

10 TAMURA, H. AND YOKOYA, N. 1984. Image database systems: A survey. Part. Recog. 17, 1, 29-43.

↑ CITINGS *36*

Kyung-Ah Han , Sung-Hyun Myaeng, Image organization and retrieval with automatically constructed feature vectors, Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval, p.157-165, August 18-22, 1996, Zurich, Switzerland

Venkat N. Gudivada , Gwang S. Jung, A linear time algorithm for retrieval by spatial constraints in multimedia database applications, Proceedings of the 1995 ACM 23rd annual conference on Computer science, p.195-202, February 28-March 02, 1995, Nashville, Tennessee, United States

Byungwoo Kim , Jong P. Yoon, Similarity measurement for aggregation of spatial objects, Proceedings of the 2005 ACM symposium on Applied computing, March 13-17, 2005, Santa Fe, New Mexico

Yung-Kuan Chan , Chih-Ya Chen, Image retrieval system based on color-complexity and color-spatial features, Journal of Systems and Software, v.71 n.1-2, p.65-70, April 2004

Amarnath Gupta , Simone Santini , Ramesh Jain, In search of information in visual media, Communications of the ACM, v.40 n.12, p.34-42, Dec. 1997

John R. Smith, Quantitative assessment of image retrieval effectiveness, Journal of the American Society for Information Science and Technology, v.52 n.11, p.969-979, September 2001

Venkat N. Gudivada , Jay Bhuyan , Ramesh Adusumilli, A retrieval technique for virtual reality databases, Proceedings of the 1997 ACM symposium on Applied computing, p.328-333, April 1997, San Jose, California, United States

Heng Tao Shen, Finding similar images quicky using object shapes, Proceedings of the tenth international conference on Information and knowledge management, October 05-10, 2001, Atlanta, Georgia, USA

Raghu Krishnapuram , Swarup Medasani , Sung-Hwan Jung , Young-Sik Choi , Rajesh

Balasubramaniam, Content-Based Image Retrieval Based on a Fuzzy Approach, IEEE Transactions on Knowledge and Data Engineering, v.16 n.10, p.1185-1199, October 2004

John R. Smith , Shih-Fu Chang, Integrated spatial and feature image query, Multimedia Systems, v.7 n.2, p.129-140, March 1999

Yeon-Jung Kim , Choon-Bo Sim , Jae-Woo Chang, Spatial match representation scheme supporting ranking in iconic images databases, Proceedings of the eighth international conference on Information and knowledge management, p.450-457, November 02-06, 1999, Kansas City, Missouri, United States

Serhan Dağtaş , Arif Ghafoor, Indexing and retrieval of video based on spatial relation sequences, Proceedings of the seventh ACM international conference on Multimedia (Part 2), p.119-122, October 30-November 05, 1999, Orlando, Florida, United States

Yuhang Wang , Fillia Makedon, R-Histogram: quantitative representation of spatial relations for similarity-based image retrieval, Proceedings of the eleventh ACM international conference on Multimedia, November 02-08, 2003, Berkeley, CA, USA

John R. Smith , Shih-Fu Chang, VisualSEEk: a fully automated content-based image query system, Proceedings of the fourth ACM international conference on Multimedia, p.87-98, November 18-22, 1996, Boston, Massachusetts, United States

Yuhang Wang , Fillia Makedon , Amit Chakrabarti, R*-Histograms: efficient representation of spatial relations between objects of arbitrary topology, Proceedings of the 12th annual ACM international conference on Multimedia, October 10-16, 2004, New York, NY, USA

Dimitris Papadias , Nikos Mamoulis , Dimitris Meretakis, Image similarity retrieval by spatial constraints, Proceedings of the seventh international conference on Information and knowledge management, p.289-296, November 02-07, 1998, Bethesda, Maryland, United States

F. C. Berger , P. Van Bommel , Th. P. Van Der Weide, Ranking Strategies for Navigation Based Query Formulation, Journal of Intelligent Information Systems, v.12 n.1, p.5-25, April 1999

Venkat N. Gudivada, TESSA—an image testbed for evaluating 2-D spatial similarity algorithms, ACM SIGIR Forum, v.28 n.2, p.17-36, Fall 1994

Dimitris Papadias, Hill climbing algorithms for content-based retrieval of similar configurations, Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval, p.240-247, July 24-28, 2000, Athens, Greece

E. Di Sciascio , M. Mongiello , F. M. Donini , L. Allegretti, Retrieval by spatial similarity: an algorithm and a comparative evaluation, Pattern Recognition Letters, v.25 n.14, p.1633-1645, 15 October 2004

S. Adali , P. Bonatti , M. L. Sapino , V. S. Subrahmanian, A multi-similarity algebra, ACM SIGMOD Record, v.27 n.2, p.402-413, June 1998

E. Di Sciascio , F. M. Donini , M. Mongiello, Spatial layout representation for query-by-sketch content-based image retrieval, Pattern Recognition Letters, v.23 n.13, p.1599-1612, November 2002

Yuhang Wang , Fillia Makedon , James Ford , Li Shen , Dina Goldin, Generating fuzzy semantic metadata describing spatial relations from images using the R-histogram, Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries, June 07-11, 2004, Tuscon, AZ, USA

Eugenio Di Sciascio , Francesco M. Donini , Marina Mongiello, A knowledge based system for content-based retrieval of Scalable Vector Graphics documents, Proceedings of the 2004 ACM symposium on Applied computing, March 14-17, 2004, Nicosia, Cyprus

Information retrieval using robust natural language processing

Full text Publisher Site Pdf (773 KB)

Source **Annual Meeting of the ACL** archive
Proceedings of the 30th annual meeting on Association for Computational Linguistics table of contents
Newark, Delaware
Pages: 104 - 111
Year of Publication: 1992

Authors Tomek Strzalkowski New York University, New York, NY
Barbara Vauthey New York University, New York, NY

Sponsor undetermined : undetermined

Publisher Association for Computational Linguistics Morristown, NJ, USA

↑ ABSTRACT

We developed a prototype information retrieval system which uses advanced natural language processing techniques to enhance the effectiveness of traditional key-word based document retrieval. The backbone of our system is a statistical retrieval engine which performs automated indexing of documents, then search and ranking in response to user queries. This core architecture is augmented with advanced natural language processing tools which are both robust and efficient. In early experiments, the augmented system has displayed capabilities that appear to make it superior to the purely statistical base.

↑ REFERENCES

Note: OCR errors may be found in this Reference List extracted from the full text article. ACM has opted to expose the complete List rather than only correct and linked references.

1 Kenneth Ward Church , Patrick Hanks, Word association norms, mutual information, and lexicography, Computational Linguistics, v.16 n.1, p.22-29, March 1990

2 W. Bruce Croft , Howard R. Turtle , David D. Lewis, The use of phrases and structured queries in information retrieval, Proceedings of the 14th annual international ACM SIGIR conference on Research and development in information retrieval, p.32-45, October 13-16, 1991, Chicago, Illinois, United States

3 C. J. Crouch, A cluster-based approach to thesaurus construction, Proceedings of the 11th annual international ACM SIGIR conference on Research and development in information retrieval, p.309-320, May 1988, Grenoble, France

4 Fagan, Joel L. 1987 Experiments in Automated Phrase Indexing for Document Retrieval: A

Comparison of Syntactic and Non-Syntactic Methods. Ph.D. Thesis, Department of Computer Science, Cornell University.

5 Ralph Grishman , Lynette Hirschman , Ngo Thanh Nhan, Discovery procedures for sublanguage selectional patterns: initial experiments, Computational Linguistics, v.12 n.3, p.205-215, July-September 1986

6 Grishman, Ralph and Tomek Strzalkowski. 1991. "Information Retrieval and Natural Language Processing." Position paper at the workshop on Future Directions in Natural Language Processing in Information Retrieval, Chicago.

7 D. Harman, Towards interactive query expansion, Proceedings of the 11th annual international ACM SIGIR conference on Research and development in information retrieval, p.321-331, May 1988, Grenoble, France

8 Harman, Donna and Gerald Candela. 1989. "Retrieving Records from a Gigabyte of text on a Minicomputer Using Statistical Ranking." Journal of the American Society for Information Science, 41 (8), pp. 581--589.

9 Harris, Zelig S. 1991. A Theory of language and Information. A Mathematical Approach. Cladendon Press. Oxford.

10 Harris, Zelig S. 1982. A Grammar of English on Mathematical Principles. Wiley.

11 Harris, Zelig S. 1968. Mathematical Structures of Language. Wiley.

12 Donald Hindle, Noun classification from predicate-argument structures, Proceedings of the 28th annual meeting on Association for Computational Linguistics, p.268-275, June 06-09, 1990, Pittsburgh, Pennsylvania

13 D. D. Lewis , W. B. Croft, Term clustering of syntactic phrases, Proceedings of the 13th annual international ACM SIGIR conference on Research and development in information retrieval, p.385-404, September 05-07, 1990, Brussels, Belgium

14 Michael L. Mauldin, Retrieval performance in Ferret a conceptual information retrieval system, Proceedings of the 14th annual international ACM SIGIR conference on Research and development in information retrieval, p.347-355, October 13-16, 1991, Chicago, Illinois, United States

15 Naomi Sager, Natural Language Information Processing: A Computer Grammmar of English and Its Applications, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 1981

16 Gerard Salton, Automatic text processing: the transformation, analysis, and retrieval of information by computer, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 1989

17 Shannon, C. E. 1948. "A mathematical theory of communication." Bell System Technical Journal, vol. 27, July-October.

18 A. F. Smeaton , C. J. van Rijsbergen, Experiments on incorporating syntactic processing of user queries into a document retrieval strategy, Proceedings of the 11th annual international ACM SIGIR conference on Research and development in information retrieval, p.31-51, May 1988, Grenoble, France

19 Sparck Jones, Karen. 1972. "Statistical interpretation of term specificity and its application in retrieval." Journal of Documentation, 28(1), pp. 11--20.

20 Sparck Jones, K. and E. O. Barber. 1971. "What makes automatic keyword classification effective?" Journal of the American Society for Information Science, May-June, pp. 166--175.

21 Sparck Jones, K. and J. I. Tait. 1984. "Automatic search term variant generation." Journal of Documentation, 40(1), pp. 50--66.

22 Tomek Strzalkowski , Barabara Vauthey, Fast text processing for information retrieval, Proceedings of the workshop on Speech and Natural Language, p.346-352, February 19-22, 1991, Pacific Grove, California

23 Strzalkowski, Tomek and Barbara Vauthey. 1991. "Natural Language Processing in Automated Information Retrieval." Proteus Project Memo #42, Courant Institute of Mathematical Science, New York University.

24 Tomek Strzalkowski, TTP: a fast and robust parser for natural language, Proceedings of the 14th conference on Computational linguistics, August 23-28, 1992, Nantes, France

25 Wilks, Yorick A., Dan Fass, Cheng-Ming Guo, James E. McDonald, Tony Plate, and Brian M. Slator. 1990. "Providing machine tractable dictionary tools." Machine Translation, 5, pp. 99--154.

↑ CITINGS *10*

Joe Zhou , Troy Tanner, Construction and visualization of key term hierarchies, Proceedings of the fifth conference on Applied natural language processing, p.307-311, March 31-April 03, 1997, Washington, DC

Tomek Strzalkowski, Natural language information retrieval: TIPSTER-2 final report, Proceedings of a workshop on held at Vienna, Virginia: May 6-8, 1996, May 06-08, 1996, Vienna, Virginia

Tomek Strzalkowski, Document representation in natural language text retrieval, Proceedings of the workshop on Human Language Technology, March 08-11, 1994, Plainsboro, NJ

Tomek Strzalkowski, TTP: a fast and robust parser for natural language, Proceedings of the 14th conference on Computational linguistics, August 23-28, 1992, Nantes, France

Tomek Strzalkowski, Building a lexical domain map from text corpora, Proceedings of the 15th conference on Computational linguistics, August 05-09, 1994, Kyoto, Japan

Alan M. Buckeridge , Richard F. E. Sutcliffe, Disambiguating noun compounds with latent semantic indexing, COLING-02 on COMPUTERM 2002: second international workshop on computational terminology, p.1-7, August 31, 2002

Suzanne Liebowitz Taylor , Deborah A. Dahl , Mark Lipshutz , Carl Weir , Lewis M. Norton , Roslyn Nilson , Marcia Linebarger, Integrated text and image understanding for document understanding, Proceedings of the workshop on Human Language Technology, March 08-11, 1994, Plainsboro, NJ

Christian Jacquemin, Recycling terms into a partial parser, Proceedings of the fourth conference on Applied natural language processing, October 13-15, 1994, Stuttgart, Germany

Marcia C. Linebarger , Lewis M. Norton , Deborah A. Dahl, A portable approach to last resort parsing and interpretation, Proceedings of the workshop on Human Language Technology, March 21-24, 1993, Princeton, New Jersey

Chengxiang Zhai, Fast statistical parsing of noun phrases for document indexing, Proceedings of the fifth conference on Applied natural language processing, p.312-319, March 31-April 03, 1997, Washington, DC

↑ Collaborative Colleagues:

Tomek Strzalkowski: Amit Bagga Ping Peng Xinyang Zhang

	Leonard Bolc	Jose Perez-Carballo
	Nick Cercone	Nobuyuki Shimizu
	Hilda Hardy	Rong Tang
	Paul B. Kantor	Liu Ting
	Fang Lin	Barabara Vauthey
	Mihnea Marinescu	Barbara Vauthey
	Miroslav Martinovic	Jin Wang
	Kwong Bor Ng	Bowden Wise
	Mark Osborn	G. Bowden Wise
Barbara Vauthey:	Tomek Strzalkowski	

↑ **Peer to Peer - Readers of this Article have also read:**

- Data structures for quadtree approximation and compression **Communications of the ACM** 28, 9
  Hanan Samet
- A hierarchical single-key-lock access control using the Chinese remainder theorem **Proceedings of the 1992 ACM/SIGAPP Symposium on Applied computing**
  Kim S. Lee , Huizhu Lu , D. D. Fisher
- The GemStone object database management system **Communications of the ACM** 34, 10
  Paul Butterworth , Allen Otis , Jacob Stein
- Putting innovation to work: adoption strategies for multimedia communication systems **Communications of the ACM** 34, 12
  Ellen Francik , Susan Ehrlich Rudman , Donna Cooper , Stephen Levine
- An intelligent component database for behavioral synthesis **Proceedings of the 27th ACM/IEEE conference on Design automation**
  Gwo-Dong Chen , Daniel D. Gajski

Efficient and effective metasearch for a large number of text databases

Full text Pdf (1.04 MB)

Source **Conference on Information and Knowledge Management** archive
Proceedings of the eighth international conference on Information and knowledge management table of contents
Kansas City, Missouri, United States
Pages: 217 - 224
Year of Publication: 1999
ISBN:1-58113-146-1

Authors
Clement Yu Dept. of EECS, University of Illinois at Chicago, Chicago, IL
Weiyi Meng Dept. of Computer Science, SUNY - Binghamton, Binghamton, NY
King-Lup Liu Dept. of EECS, University of Illinois at Chicago, Chicago, IL
Wensheng Wu Dept. of EECS, University of Illinois at Chicago, Chicago, IL
Naphtali Rishe School of Computer Science, Florida International University, Miami, FL

Sponsors SIGART: ACM Special Interest Group on Artificial Intelligence
SIGIR: ACM Special Interest Group on Information Retrieval
SIGMIS: ACM Special Interest Group on Management Information Systems

Publisher ACM Press New York, NY, USA

↑ ABSTRACT

Metasearch engines can be used to facilitate ordinary users for retrieving information from multiple local sources (text databases). In a metasearch engine, the contents of each local database is represented by a representative. Each user query is evaluated against the set of representatives of all databases in order to determine the appropriate databases to search. When the number of databases is very large, say in the order of tens of thousands or more, then a traditional metasearch engine may become inefficient as each query needs to be evaluated against too many database representatives. Furthermore, the storage requirement on the site containing the metasearch engine can be very large. In this paper, we propose to use a hierarchy of database representatives to improve the efficiency. We provide an algorithm to search the hierarchy. We show that the retrieval effectiveness of our algorithm is the same as that of evaluating the user query against all database representatives. We also show that our algorithm is efficient. In addition, we propose an alternative way of allocating representatives to sites so that the storage burden on the site containing the metasearch engine is much reduced.

↑ REFERENCES

Note: OCR errors may be found in this Reference List extracted from the full text article. ACM has opted to expose the complete List rather than only correct and linked references.

1 Christoph Baumgarten, A probabilistic model for distributed information retrieval, Proceedings of the 20th annual international ACM SIGIR conference on Research and development in information retrieval, p.258-266, July 27-31, 1997, Philadelphia, Pennsylvania, United States

2 James P. Callan , Zhihong Lu , W. Bruce Croft, Searching distributed collections with inference networks, Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval, p.21-28, July 09-13, 1995, Seattle, Washington, United States

3 Y. Fan, and S. Gauch. Adaptive Agents for Information Gathering from Multiple, Distributed Information Sources. 1999 AAAI Symposium on Intelligent Agents in Cyberspace, March 1999.

4 S. Gauch, G. Wang, and M. Gomez. ProFusion: fntelligent Fusion from Multiple, Distributed Search Engines. Journal of Universal Computer Science, 2(9), 1996.

5 Luis Gravano , Hector Garcia-Molina, Generalizing GlOSS to Vector-Space Databases and Broker Hierarchies, Proceedings of the 21th International Conference on Very Large Data Bases, p.78-89, September 11-15, 1995

6 A. Howe, and D. Dreilinger. $avvySearch: A Meta- Search Engine that Learns Which Search Engines to Query. AI Magazine, 18(2), 1997.

7 B. Kahle, and A. Medlar. An Information System for Corporate Users: Wide Area in/ormation Servers. Technical Report TMC199, Thinking Machine Corporation, April 1991.

8 T. Kirk, A. Levy, Y. Sagiv, and D. Srivastava. The information Manifold. AAAI Spring Symposium on Information Gathering in Distributed Heterogeneous Environments. 1995.

9 Martijn Koster, ALIWEB—Archie-like indexing in the WEB, Computer Networks and ISDN Systems, v.27 n.2, p.175-182, Nov. 1994

10 U. Manber, and P. Bigot. 'The Search Broker. USENIX Symposium on Internet Technologies and Systems (NSITS'97), 1997.

11 Weiyi Meng , King-Lup Liu , Clement T. Yu , Xiaodong Wang , Yuhsi Chang , Naphtali Rishe, Determining Text Databases to Search in the Internet, Proceedings of the 24rd International Conference on Very Large Data Bases, p.14-25, August 24-27, 1998

12 W. Meng, K. Liu, C. Yu, W. Wu, and N. Rishe. Estimating the Usefulness o/Search Engines. IEEE Data Engineering Conference, 1999.

13 W. Meng, C. Yu, and K. Liu. Challenges and Solutions for Building an Efficient and Effective Metaseareh Engine. Technical Report, Dept. of CS, SUNY at Binghamton, 1999.

14 Gerard Salton , Michael J. McGill, Introduction to Modern Information Retrieval, McGraw-Hill, Inc., New York, NY, 1986

15 E. Selberg, and O. Etzioni. The MetaCrawler Architecture for Resource Aggregation on the Web. IEEE Expert, 1997.

16 Ellen M. Voorhees , Narendra K. Gupta , Ben Johnson-Laird, Learning collection fusion strategies, Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval, p.172-179, July 09-13, 1995, Seattle, Washington, United States

17 Jinxi Xu , Jamie Callan, Effective retrieval with distributed collections, Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval, p.112-120, August 24-28, 1998, Melbourne, Australia

18 Clement Yu , King-Lup Liu , Wensheng Wu , Weiyi Meng , Naphtali Rishe, Finding the Most Similar Documents across Multiple Text Databases, Proceedings of the IEEE Forum on Research and Technology Advances in Digital Libraries, p.150, March 19-21, 1999

19 C. Yu, K. Liu, W. Wu, W. Meng and N. Rishe. A Methodology to Retrieve Text Documents from Multiple Databases. Tech. report, U. of Illinois at Chicago, 1999.

20 Clement T. Yu , Weiyi Meng, Principles of database query processing for advanced applications, Morgan Kaufmann Publishers Inc., San Francisco, CA, 1998

21 C. Yu, W. Meng, K. Liu, W. Wu, and N. Rishe. Efficient and Effective Metasearch for a Large Num. ber of Tezt Databases. Tech. report, U. of Illinois at Chicago, 1999. (http://panda.cs.binghamton.edu/ ~meng/cikm99.ps)

22 Budi Yuwono , Dik Lun Lee, Server Ranking for Distributed Text Retrieval Systems on the Internet, Proceedings of the Fifth International Conference on Database Systems for Advanced Applications (DASFAA), p.41-50, April 01-04, 1997

↑ CITINGS *16*

King-Lup Liu , Adrain Santoso , Clement Yu , Weiyi Meng, Discovering the representative of a search engine, Proceedings of the tenth international conference on Information and knowledge management, October 05-10, 2001, Atlanta, Georgia, USA

King-Lup Liu , Clement Yu , Weiyi Meng, Discovering the representative of a search engine, Proceedings of the eleventh international conference on Information and knowledge management, November 04-09, 2002, McLean, Virginia, USA

Robert W. P. Luk , Tharam S. Dillon , Vincent T. Y. Ng, Supporting metasearch with XSL, Journal of Systems and Software, v.73 n.1, p.159-168, September 2004

Shengli Wu , Fabio Crestani, Shadow document methods of resutls merging, Proceedings of the 2004 ACM symposium on Applied computing, March 14-17, 2004, Nicosia, Cyprus

Yugyung Lee , Changgyu Oh , Eun Kyo Park, Intelligent knowledge discovery in peer-to-peer file sharing, Proceedings of the eleventh international conference on Information and knowledge management, November 04-09, 2002, McLean, Virginia, USA

Matthias Bender , Sebastian Michel , Peter Triantafillou , Gerhard Weikum , Christian Zimmer, MINERVA: collaborative P2P search, Proceedings of the 31st international conference on Very large data bases, August 30-September 02, 2005, Trondheim, Norway

B. Uygar Oztekin , George Karypis , Vipin Kumar, Expert agreement and content based reranking in a meta search environment using Mearf, Proceedings of the 11th international conference on World Wide Web, May 07-11, 2002, Honolulu, Hawaii, USA

James C. French , Allison L. Powell , Fredric Gey , Natalia Perelman, Exploiting a controlled vocabulary to improve collection selection and retrieval effectiveness, Proceedings of the tenth international conference on Information and knowledge management, October 05-10, 2001, Atlanta, Georgia, USA

Panagiotis G. Ipeirotis , Luis Gravano, When one sample is not enough: improving text database selection using shrinkage, Proceedings of the 2004 ACM SIGMOD international conference on Management of data, June 13-18, 2004, Paris, France

Clement Yu , Weiyi Meng , Wensheng Wu , King-Lup Liu, Efficient and effective metasearch for text databases incorporating linkages among documents, ACM SIGMOD Record, v.30 n.2, p.187-198, June

2001

Zonghuan Wu , Weiyi Meng , Clement Yu , Zhuogang Li, Towards a highly-scalable and effective metasearch engine, Proceedings of the tenth international conference on World Wide Web, p.386-395, May 01-05, 2001, Hong Kong, Hong Kong

Clement Yu , King-Lup Liu , Weiyi Meng , Zonghuan Wu , Naphtali Rishe, A Methodology to Retrieve Text Documents from Multiple Databases, IEEE Transactions on Knowledge and Data Engineering, v.14 n.6, p.1347-1361, November 2002

Weiyi Meng , Zonghuan Wu , Clement Yu , Zhuogang Li, A highly scalable and effective method for metasearch, ACM Transactions on Information Systems (TOIS), v.19 n.3, p.310-335, July 2001

Allison L. Powell , James C. French, Comparing the performance of collection selection algorithms, ACM Transactions on Information Systems (TOIS), v.21 n.4, p.412-456, October 2003

Weiyi Meng , Clement Yu , King-Lup Liu, Building efficient and effective metasearch engines, ACM Computing Surveys (CSUR), v.34 n.1, p.48-89, March 2002

Robert W.P. Luk , H. V. Leong , Tharam S. Dillon , Alvin T.S. Chan , W. Bruce Croft , James Allan, A survey in indexing and searching XML documents, Journal of the American Society for Information Science and Technology, v.53 n.6, p.415-437, May, 2002

↑ INDEX TERMS

Primary Classification:

H. Information Systems

↳ **H.3** INFORMATION STORAGE AND RETRIEVAL

↳ **H.3.3** Information Search and Retrieval

↳ **Subjects:** Search process

Additional Classification:

H. Information Systems

↳ **H.2** DATABASE MANAGEMENT

↳ **H.3** INFORMATION STORAGE AND RETRIEVAL

↳ **H.3.3** Information Search and Retrieval

↳ **Subjects:** Retrieval models

↳ **H.3.5** On-line Information Services

↳ **Subjects:** Web-based services

↳ **H.5** INFORMATION INTERFACES AND PRESENTATION (I.7)

↳ **H.5.3** Group and Organization Interfaces

↳ **Subjects:** Web-based interaction

General Terms:

Algorithms, Design, Documentation, Experimentation, Management, Measurement, Performance, Theory

↑ Collaborative Colleagues:

King-Lup Liu:	Yuhsi Chang
	Weiyi Meng

	Naphtali Rishe Adrain Santoso A. Prasad Sistla Xiaodong Wang Wensheng Wu Zonghuan Wu Clement Yu Clement T. Yu			
Weiyi Meng:	Sudhir Aggarwal Komanduru Sai C Yuhsi Chang Rongquen Chen Yuxi Chen Son Dao AnHai Doan Chun Du Ophir Frieder David Grossman Hai He Inje Kim	Won Kim Elizabeth Laun Zhuogang Li Chengwen Liu Fang Liu K. L. Liu King-Kup Liu King-Lup Liu Shuang Liu Qian Peng Tracy Pham Hua Qian	Yan Qin Vijay Raghavan Vuyyuru Rama N. Rishe Naphtali Rishe Adrain Santoso Prasoon Sharma Wei Sun Gaoming Wang Wei Wang Wenxian Wang Xiaodong Wang	Wensheng Wu Z. Wu Zonghuan Wu C. Yu Clement Yu Clement T. Yu Changqing Zhang Weining Zhang Yi Zhang Zhongfei Zhang Hongkun Zhao
Naphtali Rishe:	Michael Alexopoulos Y. Alp Aslandogan Rukshan Athauda Rukshan Indika Athauda David Barton Yaman Battikhi Toby Berk Dmitriy A. Beryoza Guohong Cao Hu Cao Sam Chamberlain Yuhsi Chang Maxim Chekmasov Marina Chekmasova Chung-Min Chen Shu-Ching Chen Son Dao Ian De Felipe Minglin Deng Yi Deng Oksana Dyganova Raimund K. Ege Scott Graham	Scott C. Graham Martha E. Gutierrez Liqin Jiang Vladimir Lebedev Qiang Li Hai Lin Yibei Ling G. J. Lipovski Chengwen Liu Jingyu Liu K. L. Liu King-Lup Liu Leonardo Loureiro Leonardo Haddad Loureiro Xiaoling Lu Xiaobin Ma Daniel Mendez Alejandro Mendoza Weiyi Meng Khaled Samih Naboulsi Sham Navathe Carlos Ordonez Cyril Orji	Aris Ouksel Luis Pachas Philippe Pardo Bernard C. Parenteau Ginanina G. Rocha Rosany Hernandez Rodriguez Mario Sanchez Mario R. Sanchez Andriy Selivonenko Artyom Shaposhnikov Mukesh Singhal A. Prasad Sistla Prasad Sistla Wei Sun Yanli Sun D. Tal Doron Tal Chuck Thier Goce Trajcevski Jinny Uppal Alexander Vaschillo Dmitry Vasilevsky Wei Wang	Xiaodong Wang Xinran Wang Mark Allen Weiss Ouri Wolfson Jing Wu Wensheng Wu Zonghuan Wu Bo Xu Li Yang Qi Yang Xiangyu Ye Yelena Yesha C. Yu Clement Yu Clement T. Yu Jun Yuan Fengli Zhang Jutai Zhou O. D. Zhukov Jon Zou
Wensheng Wu:	AnHai Doan King-Lup Liu Weiyi Meng Naphtali Rishe Clement Yu			
Clement Yu:	Y. Alp Aslandogan Dina Bitton Richard Bruno Komanduru Sai C Isabel Cruz Carlo A. Curino Son Dao AnHai Doan	Ashfaq Khokhar Won Kim Franky Kin-Pong Chan Bruce Lambert Bruce L. Lambert Zhuogang Li Yibei Ling	Weiyi Meng Aris Ouksel Qian Peng Hua Qian Yan Qin Vijay Raghavan Vuyyuru Rama Naphtali Rishe	Chuck Thier Mohanraj Thirumalai John Tullis Ada Wai-chee Fu Wei Wang Wenxian Wang

Chun Du
Ophir Frieder
David Grossman
Hai He
Biao Jia
Yuanyuan Jia
Bing Liu
Chengwen Liu
Fang Liu
King-Kup Liu
King-Lup Liu
Shuang Liu
Chengjie Luo
Adrain Santoso
Peter Scheuermann
Prasoon Sharma
A. Prasad Sistla
Prasad Sistla
Wei Sun
Jane Wells
Patricia M. West
Ouri Wolfson
Jing Wu
Wensheng Wu
Zonghuan Wu
Qi Yang
Hongkun Zhao